

Diagnostic Accuracy of Digital Solutions for Screening for Cognitive Impairment: A Systematic Review and Meta-Analysis

Marisa Magno ^{1,*}, Ana Isabel Martins ², Joana Pais ^{3,4}, Anabela G. Silva ² and Nelson Pacheco Rocha ⁵¹ Department of Medical Sciences, University of Aveiro, 3810-193 Aveiro, Portugal² Center for Health Technology and Services Research—CINTESIS@RISE, School of Health Sciences, University of Aveiro, 3810-193 Aveiro, Portugal; anaisabelmartins@ua.pt (A.I.M.); asilva@ua.pt (A.G.S.)³ EPIUnit—Institute of Public Health, Laboratory for Integrative and Translational Research in Population Health (ITR), University of Porto, 4050-600 Porto, Portugal; joanapais@neuroinova.com⁴ Neuroinova, Lda., 4410-463 Vila Nova de Gaia, Portugal⁵ IEETA—Institute of Electronics and Informatics Engineering of Aveiro, Department of Medical Sciences, University of Aveiro, 3810-193 Aveiro, Portugal; npr@ua.pt

* Correspondence: marisamagno@ua.pt

Abstract: The early detection of cognitive impairment is essential in order to initiate interventions and guarantee access to healthcare services. Digital solutions are emerging in the literature as an alternative approach to cognitive screening. Our primary goal is to synthesize the evidence on digital solutions' diagnostic ability to screen for cognitive impairment and their accuracy. A secondary goal is to distinguish whether the ability to screen for cognitive impairment varies as a function of the type of digital solution: paper-based or innovative digital solutions. A systematic review and meta-analysis of digital solutions' diagnostic accuracy were conducted, including 25 studies. Digital solutions presented a variable diagnostic accuracy range. Innovative digital solutions offered at least 0.78 of sensitivity but showed lower specificity levels than the other subgroup. Paper-based digital solutions revealed at least 0.72 of specificity, but sensitivity started at 0.49. Most digital solutions do not demand the presence of a trained professional and include an automatic digital screening system and scoring, which can enhance cognitive screening and monitoring. Digital solutions can potentially be used for cognitive screening in the community and clinical practice, but more investigation is needed for an evidence-based decision. A careful assessment of the accuracy levels and quality of evidence of each digital solution is recommended.

Keywords: cognitive screening; cognitive impairment; diagnostic accuracy; digital solutions

Citation: Magno, M.; Martins, A.I.; Pais, J.; Silva, A.G.; Rocha, N.P. Diagnostic Accuracy of Digital Solutions for Screening for Cognitive Impairment: A Systematic Review and Meta-Analysis. *Appl. Sci.* **2024**, *14*, 2640. <https://doi.org/10.3390/app14062640>

Academic Editor: Qi-Huang Zheng

Received: 29 February 2024

Revised: 8 March 2024

Accepted: 19 March 2024

Published: 21 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

There are several known risk factors (e.g., diabetes, hypertension, hypercholesterolemia, depression, physical frailty, a low education level, or a low social support level) contributing to neurodegenerative diseases such as Alzheimer, Parkinson, Huntington, or frontotemporal dementia, but aging is the strongest one [1–6]. Therefore, the prevalence of these diseases increases as our society ages.

Mild cognitive impairment (MCI), an intermediate stage between normal aging and dementia, is characterized by an objective cognitive decline in one or more cognitive domains (e.g., memory, attention, language, or executive function) without any significant impairment in daily-life activities [7], and may be associated with a variety of underlying causes, including Alzheimer's pathophysiology [7–9]. In turn, dementia is a major neurocognitive disorder that is characterized by a significant decline in one or more cognitive domains that interferes with a person's independence in daily activities [10]. Although there is evidence that patients with MCI may experience reversion to cognitive normality [10,11], there is a high probability that this condition will progress to dementia. Therefore, early detection of

MCI is critical to effectively initiate the intervention (including counseling, psychoeducation, cognitive training, and medication [12]), and guarantee to both patients and relatives access to relevant healthcare services [13]. However, MCI is significantly misdiagnosed due to a diverse set of barriers, namely the high prevalence of comorbidities among older adults, a lack of expertise or limited confidence of the practitioners, the short duration of most primary care visits, limitations of the assessment instruments, or the inadequacy of electronic health record systems in terms of the integration of cognitive assessments, which limits the ability to track an individual's cognitive function over time [7].

Despite these barriers, there are many screening tests that provide a quick evaluation of cognitive and functional aspects. At present, two of the most well-known cognitive screening tests are the Mini-Mental State Examination (MMSE) [14] and the Montreal Cognitive Assessment (MoCA) [15], which include tasks for assessing multiple cognitive domains. In addition to MMSE and MoCA, other currently available cognitive tests also encompass multiple cognitive domains, including the Neuropsychiatry Unit Cognitive Assessment Tool (NUCOG) [16], the Saint Louis University Mental Status examination (SLUMS) [17], the Self-Administered Gerocognitive Examination (SAGE) [18], or Addenbrooke's Cognitive Examination III (ACE-III) [19]. In turn, screening tests such as the Alzheimer Quick Test (AQT) [20], Scenery Picture Memory Test (SPMT) [21], Memory Impairment Screen (MIS) [22], Mini-Cog [23,24], or Clock Drawing [25] measure one or two cognitive domain (i.e., attention for the AQT, episodic memory for the SPMT, memory and orientation for the MIS, memory and visuospatial abilities for the Mini-Cog, or executive functions and visuospatial abilities for the Clock Drawing), but require less than five minutes to be applied [26].

Computerized solutions to support neuropsychological tests have existed for several decades and might use different types of interaction devices, be it computers, handheld devices, or virtual reality [27]. Some solutions offer adaptations of paper-based tests to evaluate specific cognitive domains [28] (e.g., the Trail-Making Test or Simple and Complex Reaction Time) or multiple cognitive domains [27] (e.g., MoCA [29], MMSE [30], or SAGE [31]), while other solutions (e.g., Memoro [32], the NutriNet-Santé Cognitive Test Battery (NutriCog) [33], or the Cambridge Neuropsychological Test Automated Battery (CANTAB) [34]) were specifically developed to be applied using electronic means.

In recent years, several innovative solutions have been developed for diagnosing, monitoring (e.g., artificial intelligence applied to radiomics analysis [35]), and managing cognitive impairment (e.g., digital solutions to support self-management in older people with cognitive impairment [36]). Furthermore, the scientific literature reports the development of new instruments that are able to monitor individuals in their residential environments without the presence of a health professional [27]. This possibility maximizes flexibility and widens people's access to cognitive assessment at lower costs [27]. In this respect, smart devices (e.g., smartphones, smartwatches, or smart-home devices) may collect data on individuals' habits and patterns, which can be analyzed to detect subtle changes that may indicate a decline in cognitive performance [37]. Moreover, serious games and virtual reality are alternative approaches to cognitive screening and may also reduce feelings of test anxiety [37,38].

The research question to be addressed in this systematic review and meta-analysis is: how accurate are digital solutions for detecting both the presence and absence of cognitive impairment in individuals aged 18 years old and over? The primary goal of the current review is to synthesize the evidence on digital solutions' diagnostic ability to screen for cognitive impairment and their accuracy. A secondary goal is to distinguish whether the ability to screen cognitive impairment varies as a function of the type of digital solution: (1) based, in essence, on pre-existing traditional tests, named paper-and-pencil tests (abbreviated as paper-based digital solutions throughout this article); (2) developed from their inception to be applied by electronic means (abbreviated as innovative digital solutions throughout this article).

2. Materials and Methods

2.1. Protocol Registration

This systematic review was conducted considering the recommendations of the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy [39], and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [40]. The protocol was registered in PROSPERO [41] on the 2nd of November 2021 (CRD42021282993).

2.2. Search Strategy and Study Eligibility Criteria

The search was performed in Scopus, Web of Science, and PubMed in September 2022. Databases were searched from inception to August 2022 using the following Boolean expression: ('cognitive screening' OR 'cognitive test' OR 'memory screening' OR 'memory test' OR 'attention screening' OR 'attention test') AND ('computer' OR 'game' OR 'gaming' OR 'virtual' OR 'online' OR 'internet' OR 'mobile' OR 'app' OR 'digital').

To be included in this review, studies had to: (i) focus on any digital solution (the index test, i.e., the new or alternative test whose accuracy is being evaluated against a reference standard, i.e., the test against which the index test is being compared and that is considered a "gold standard" [42]) that can be used as a generic community-based screening tool for cognitive impairment, and that was self-administrated, i.e., performed independently by the participant without a professional conducting the test [27]; (ii) include a sample of adults (≥ 18 years old) or older adults (≥ 65 years old); (iii) compare the digital solution with a reference standard (i.e., another instrument, a clinical assessment, or a combination of these); (iv) be written in English; (v) follow case-control, cross-sectional, or cohort designs that at some point allow for the identification of two groups (with and without cognitive impairment); (vi) report at least one diagnostic accuracy property, namely sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), or, alternatively, provide enough data to calculate these indicators. Studies that included participants that had any acute neurological condition or cognitive impairment, or were institutionalized were excluded. In addition, studies that reported on digital solutions used as a monitoring tool for patients with an existing cognitive impairment diagnosis were also excluded.

2.3. Study Selection Process and Data Extraction

All retrieved references were imported into the Mendeley Desktop software, Version 1.19.8, and checked for duplicates by one author (NPR). This author (NPR) screened the titles and abstracts of all citations according to the predefined study-selection criteria. Then, the full texts of potentially relevant articles were retrieved and independently assessed by two randomly chosen authors from a set of three authors (AGS, AIM, and NPR), to verify if the inclusion and exclusion criteria were met. If a consensus could not be reached between the two authors, the third author was consulted.

Data from included studies were extracted by two authors (AIM and MM) using an electronic form developed for this purpose. The extracted information was revised and discussed with the other two authors (AGS and NPR). The information extracted from each study was: the author(s) and year of publication; the sample sizes and characteristics (e.g., sex, age); the type and name of the digital solution (index test); the type and name of the reference standard test; and the diagnostic accuracy property (e.g., estimates of sensitivity and specificity). For each study, the information used to construct a two-by-two contingency table for each index test, including the number of True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) results was also extracted. When these counts were missing from the study, the data needed (e.g., sample size, number of participants with the target condition, estimates of sensitivity and specificity, and estimates of PPV and NPV) were extracted.

The results presented in this review consider the best cut-off reported for each index test in each study for achieving the best diagnostic ability to screen for cognitive impairment.

If more than one index test result was presented (e.g., different thresholds), we chose the results given by the better cut-off reported, considering the reference standard test. Sensitivity and specificity depend on the cut-off value considered positive to identify the target condition (i.e., generally, the higher the sensitivity, the lower the specificity, and the higher the specificity, the lower the sensitivity) [43,44].

2.4. Methodological Quality Assessment

Each manuscript was independently assessed by two randomly chosen authors from a set of three authors (AGS, AIM, and NPR). Disagreements were solved by consensus or discussion with the third author. The assessment of the eligible studies' methodological quality was performed using the Quality Assessment of Diagnostic Accuracy Studies-2 tool (QUADAS-2). QUADAS-2 is a validated tool used to evaluate the quality of diagnostic accuracy studies [45], and comprises four domains: patient selection, index test, reference standard, and flow and timing. Each domain is assessed in terms of the risk of bias through signaling questions that can be answered with "yes", "no", or "unclear". The first three domains (i.e., patient selection, index test, and reference standard) are also assessed based on applicability concerns. Overall concerns about the risk of bias and applicability per domain are then rated as "high", "low", or "unclear". These results are defined for each domain [45]. A pilot test for the bias risk assessment was conducted, using studies that were not eligible for this review.

2.5. Quality of the Evidence

The overall quality (certainty) of the evidence for each meta-analysis was assessed independently by two authors (MM and AGS), according to the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach [46,47]. The GRADE approach guided the assessment and the rating of the evidence's quality and confidence considering the domains of risk of bias, inconsistency, indirectness, and publication bias. For the publication bias assessment, additional statistical analyses were conducted, namely the test for funnel plot asymmetry (Deek's test). The quality of the evidence was rated, based on the assessment of each domain, as "high", "moderate", "low", or "very low".

2.6. Data Analysis

For each study, a two-by-two contingency table was constructed, including the TP, FP, TN, and FN for the index tests. If these values were not reported in the manuscript, they were calculated from the data extracted from each study (sample size, number of participants with the target condition, estimates of sensitivity and specificity, or estimates of PPV and NPV), following the recommendations of the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy for the calculation of TP, FP, TN, and FN results [48]. Approximations and rounding were made, if necessary. Calculations were double-checked and cross-checked against the accuracy measures presented in the study.

For the meta-analysis, we used hierarchical random-effects models and Receiver Operating Characteristic (ROC) analysis. Hierarchical Summary Receiver Operating Characteristics (HSROC) models were implemented for the estimation of a Summary Receiver Operating Characteristic (SROC) curve. This method provides information on the test performance, describing variations in sensitivity and specificity [43,49], considering the recommendations of the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy [39].

A SROC plot was developed, presenting the results of each study in the ROC space, the Summary ROC (SROC) curve, the summary estimates of sensitivity, and the false positive rate (1—specificity), with 95% confidence and 95% predictive regions.

A sensitivity analysis was performed by removing the studies of paper-based digital solutions and displaying the Summary ROC (SROC) curves, summary estimates of sensitivity, and the false positive rate (1—specificity), with 95% confidence and 95% predictive regions, in a SROC plot.

Data were subdivided into two subgroups considering the type of index test used: (i) paper-based digital solutions; (ii) innovative digital solutions. A SROC plot was also developed for each index test subgroup. The estimate of the summary points and confidence intervals (CI) for sensitivity and specificity were calculated.

To perform the meta-analysis, a web application developed in R (R Core Team, Vienna) using Shiny, the MetaDTA, was used [50,51]. Among other features, MetaDTA allows to incorporate the data obtained by the QUADAS-2 tool into the graphical representation.

3. Results

3.1. Study Selection

The results of the search performed on databases are presented in Figure 1. A total of 8557 articles were identified. In the first step, 3452 duplicate articles, 311 reviews or surveys, 171 references without an abstract or without authors, 141 articles not written in English, and one article retracted, were removed. After that, 4481 articles remained for screening based on the title and abstract. Of these, 4373 articles were excluded because they did not meet the outlined inclusion criteria, whereas 108 full-text articles were thought potentially eligible. Twenty-five studies were included in this systematic review according to the eligibility criteria (Figure 1).

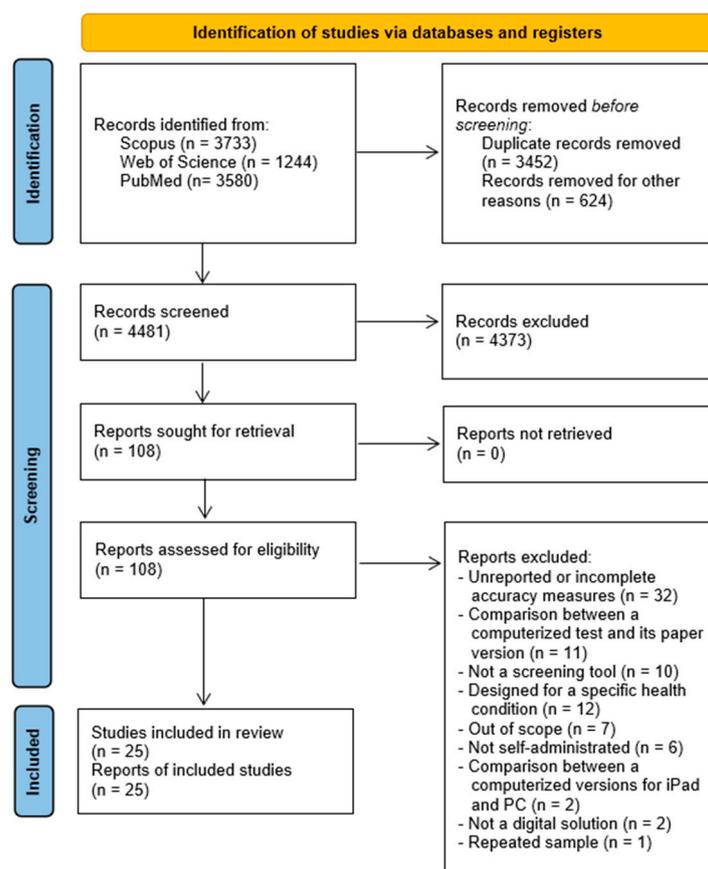


Figure 1. PRISMA flow diagram.

3.2. Methodological Quality

The results of the QUADAS-2 assessment are summarized in Table 1 and displayed in Figure 2. The risk of bias in the flow and timing domain is low in 17 out of the 25 diagnostic accuracy studies evaluated. The risk of bias in the reference standard and index text domains is unclear in 16 out of the 25 studies. Eighteen studies present a high risk of bias in the patient-selection domain. Concerns about risk applicability for the patient-selection domain were rated as high in 12 studies, low in 12 studies, and unclear in 1 study. Concerns

on the applicability in the domains of the reference standard and the index test were rated as low for most of the studies. The exceptions were three studies [52–54] that scored high in the domain of the reference standard and one study [55] that scored high in the domain of the index test.

Table 1. QUADAS-2 assessment results—tabular display.

Study	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Curriel et al. (2016) [56]	⊖	?	⊖	⊖	⊖	⊖	⊖
Alegret et al. (2020) [57]	⊖	⊖	?	⊖	⊖	⊖	⊖
Kokubo et al. (2018) [52]	⊖	?	?	⊖	⊖	⊖	⊖
Rhodium-Meester et al. (2020) [53]	⊖	⊖	?	?	⊖	⊖	⊖
Ichii et al. (2019) [58]	⊖	?	?	?	⊖	⊖	⊖
Wong et al. (2017) [59]	⊖	?	?	⊖	⊖	⊖	⊖
Buckley et al. (2017) [60]	⊖	?	?	⊖	⊖	⊖	⊖
Memória et al. (2014) [61]	⊖	?	?	⊖	⊖	⊖	⊖
Ke Yu et al. (2015) [62]	⊖	⊖	⊖	⊖	⊖	⊖	⊖
Scharre et al. (2017) [63]	⊖	⊖	⊖	⊖	⊖	⊖	⊖
Tierney et al. (2014) [64]	⊖	?	?	⊖	⊖	⊖	⊖
Brandt et al. (2014) [54]	⊖	?	⊖	⊖	⊖	⊖	⊖
Van Mierlo et al. (2017) [65]	⊖	⊖	⊖	?	⊖	⊖	⊖
Dougherty Jr. et al. (2010) [66]	⊖	⊖	⊖	?	⊖	⊖	⊖
Van der Hoek et al. (2019) [67]	⊖	?	?	?	⊖	⊖	⊖
Scanlon et al. (2015) [68]	⊖	⊖	⊖	⊖	⊖	⊖	⊖
Fung et al. (2020) [69]	⊖	?	?	⊖	⊖	⊖	⊖
Cahn-Hidalgo et al. (2020) [70]	⊖	?	?	⊖	⊖	⊖	⊖
Kalafatis et al. (2021) [55]	⊖	?	?	⊖	⊖	⊖	⊖
Rodríguez-Salgado et al. (2021) [71]	⊖	?	?	⊖	⊖	⊖	⊖
Liu et al. (2021) [72]	⊖	⊖	⊖	⊖	⊖	⊖	⊖
Yan et al. (2021) [73]	⊖	?	?	⊖	⊖	⊖	⊖
Ye et al. (2022) [74]	⊖	?	?	⊖	⊖	⊖	⊖
Paterson et al. (2022) [75]	⊖	⊖	⊖	⊖	?	⊖	⊖
Cheah et al. (2022) [76]	⊖	?	?	⊖	⊖	⊖	⊖

⊖ : low risk; ⊖ : high risk; ? : unclear risk.

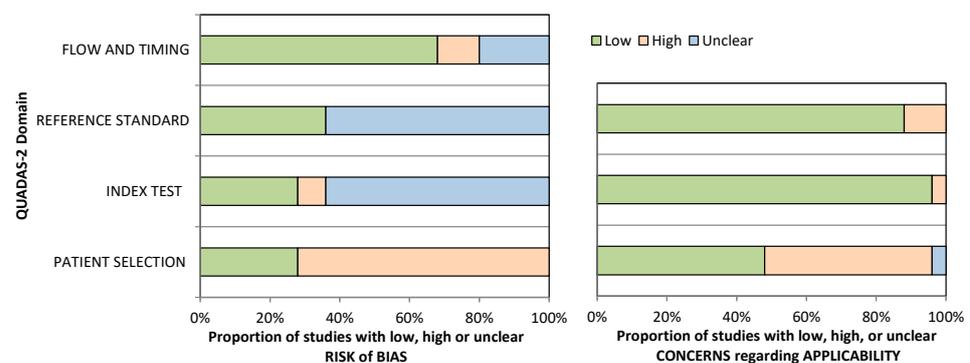


Figure 2. QUADAS-2 assessment results—graphical display.

3.3. General Overview of Included Studies

The studies included in this review adopted distinct definitions of the target condition. Most of them (18 out of the 25) defined the target condition as Mild Cognitive Impairment (MCI), including two studies that used different terminologies, namely Subtle Cognitive Impairment (SCI) and Mild Cognitive Dementia (MCD). One study considered amnesic Mild Cognitive Impairment (aMCI) as the target condition. Three studies included MCI or Mild Impairment (MI) and other clinical conditions (e.g., MCI and Dementia, MCI and Mild Alzheimer's Disease, MI and Impairment). Three studies specified Cognitive Impairment (CI) as the target condition; one of them included a significant percentage of severe cases of dementia and was excluded from the meta-analysis [66], as the other studies did not have a substantial proportion of severe cases of dementia in their samples.

Twenty-three studies out of the 25 studies included used distinct instruments and/or clinical assessment processes as reference standards.

The index tests differed across all included studies, except for two studies that used the MemTrax test (MTX) [67,72], and two studies that used the Brain Health Assessment test (BHA) [71,75]. Subgrouping the studies, there were 16 studies reporting on paper-based digital solutions and 9 studies reporting on innovative digital solutions. The index tests reporting on paper-based digital solutions varied from a direct transposition of the original paper-based test [62,63] to a substantial modification, including visual and visuospatial tasks [65] or a creation of a virtual environment [64]. Regarding the innovative digital solutions, this is also a very diverse subgroup, including solutions involving digital tasks [58,59], virtual reality and gaming [73], and artificial intelligence methods [55,76]. The characteristics of the included studies, the reference standard used, and the index tests' description can be found in the Supplementary Materials.

3.4. Meta-Analysis Results

The meta-analysis included 24 studies. The HSROC models for the estimation of a Summary ROC (SROC) curve project the results of each of the 24 studies in the ROC space, with the covariate of the index test subgroup, and display the summary estimates of sensitivity and the false positive rate (1—specificity), with 95% confidence and 95% predictive regions (Figure 3). The results indicate that sensitivity values for the index tests vary between 0.49 and 0.95 and the specificity values vary between 0.50 and 0.91. Innovative digital solutions presented values for sensitivity that vary between 0.78 and 0.94 and specificity values that vary between 0.50 and 0.90. The sensitivity of paper-based digital solutions varies between 0.49 and 0.95 and their specificity varies between 0.72 and 0.91.

The results of the sensitivity analysis of the original model (a random-effects meta-analysis of all digital solutions), when removing the studies of paper-based digital solutions (sensitivity analysis model), are displayed in Figure 4.

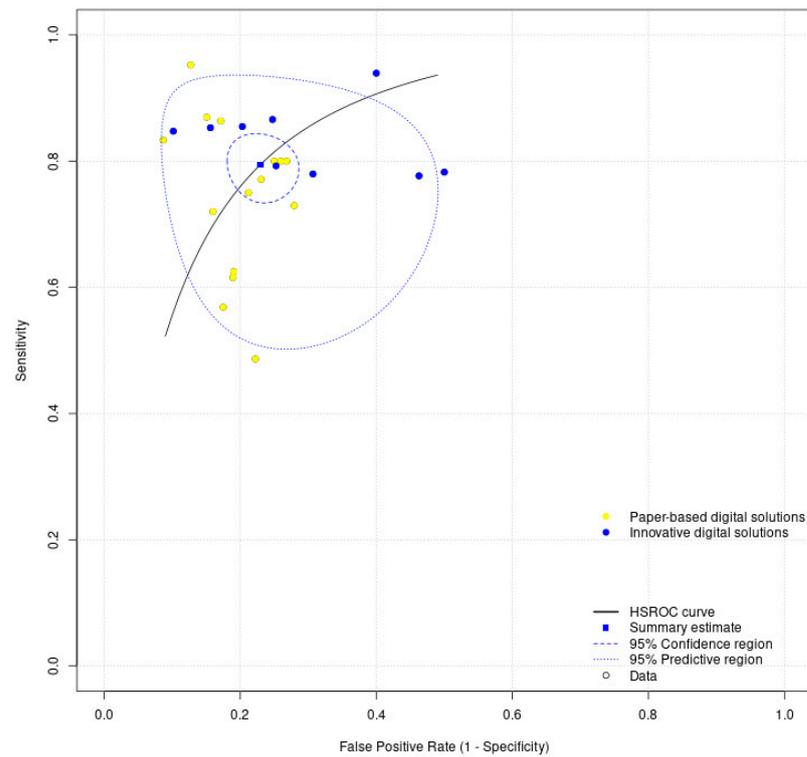


Figure 3. Random-effects meta-analysis of all digital solutions—Summary ROC (SROC) curve, summary estimates of sensitivity and the false positive rate (1—specificity), with 95% confidence and 95% predictive regions, presenting the covariate index test subgroup.

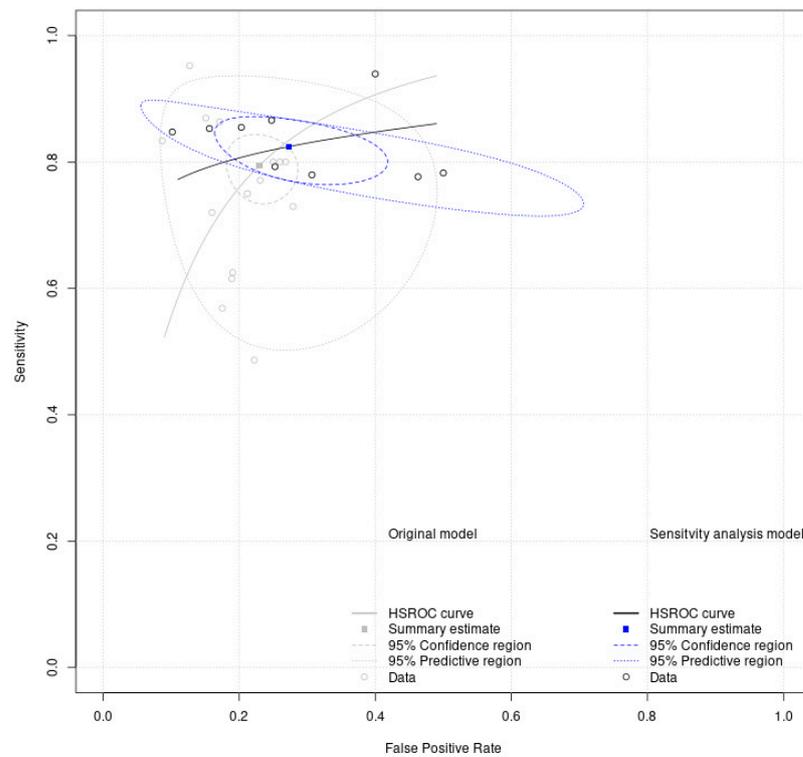
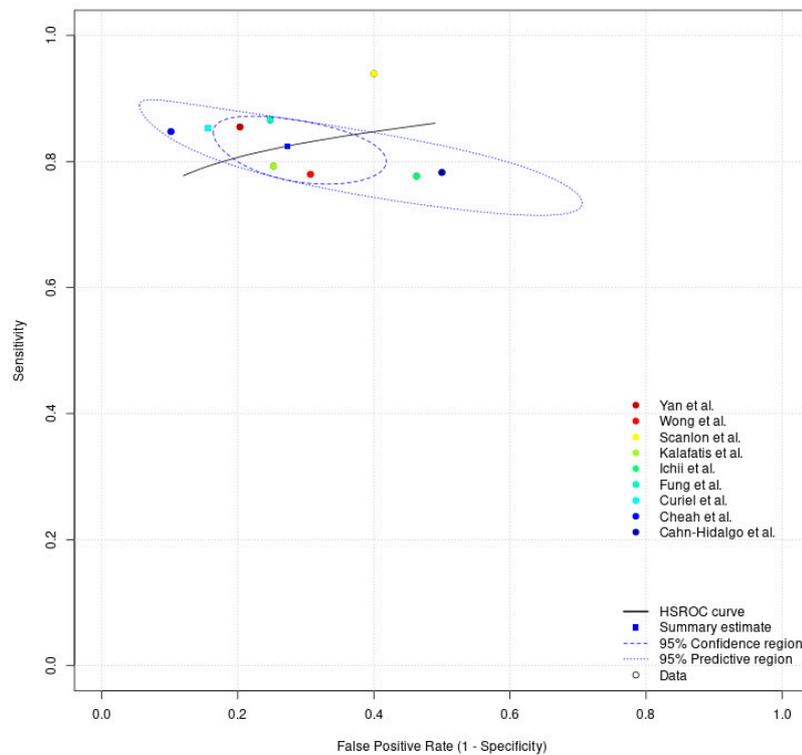


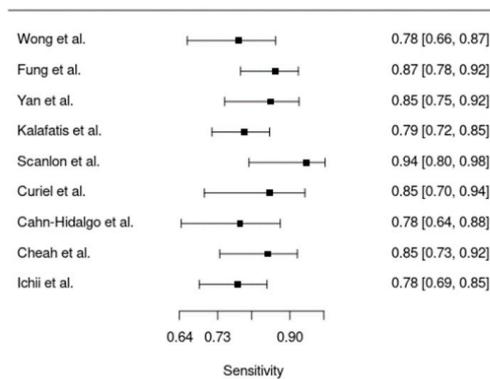
Figure 4. Original model (random-effects meta-analysis of all digital solutions) sensitivity analysis, without the studies of paper-based digital solutions (sensitivity analysis model)—Summary ROC (SROC) curve, summary estimates of sensitivity and the false positive rate (1—specificity), with 95% confidence and 95% predictive regions.

Figures 5 and 6 display the summary estimates of sensitivity and the false positive rate (1—specificity), presenting the 95% confidence and 95% predictive regions for the index tests subgrouped according to the type of test (paper-based digital solutions or innovative digital solutions). Each study is represented with a circle in the ROC space. The forest plots of sensitivity and specificity are presented per study.



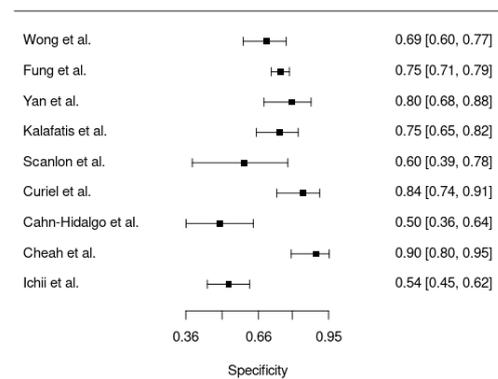
(a)

Forest plot of sensitivity



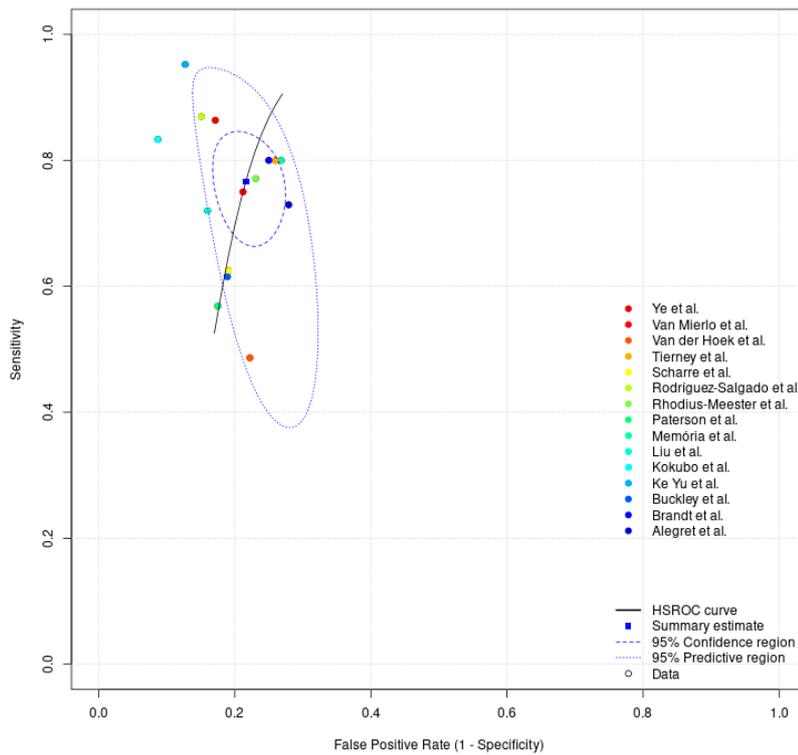
(b)

Forest plot of specificity



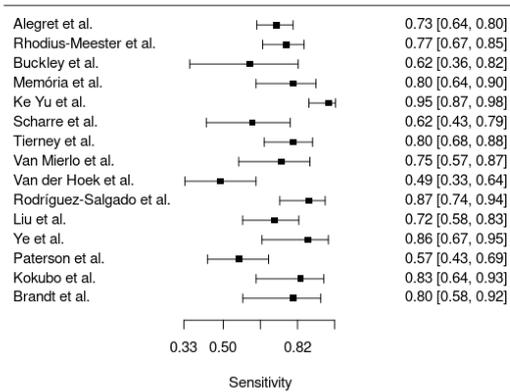
(c)

Figure 5. (a) Random-effects meta-analysis of innovative digital solutions—Summary ROC (SROC) curve, summary estimates of sensitivity and the false positive rate (1—specificity), with 95% confidence and 95% predictive regions; (b) forest plot of sensitivity; (c) forest plot of specificity. References: Cahn-Hidalgo et al. [70], Cheah et al. [76], Curiel et al. [56], Fung et al. [69], Ichii et al. [58], Kalafatis et al. [55], Scanlon et al. [68], Yan et al. [73], Wong et al. [59].



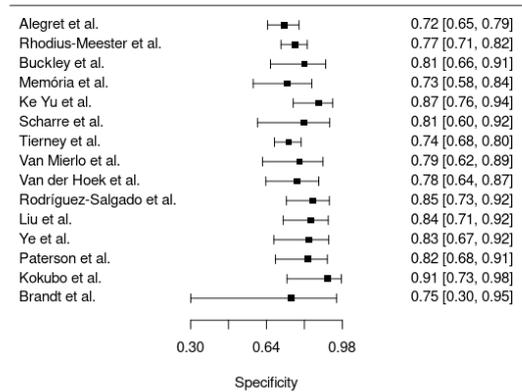
(a)

Forest plot of sensitivity



(b)

Forest plot of specificity



(c)

Figure 6. (a) Random-effects meta-analysis of paper-based digital solutions—Summary ROC (SROC) curve, summary estimates of sensitivity and the false positive rate (1—specificity), with 95% confidence and 95% predictive regions; (b) forest plot of sensitivity; (c) forest plot of specificity. References: Alegret et al. [57], Brandt et al. [54], Buckley et al. [60], Ke Yu et al. [62], Kokubo et al. [52], Liu et al. [72], Memória et al. [61], Paterson et al. [75], Rhodius-Meester et al. [53], Rodríguez-Salgado et al. [71], Scharre et al. [63], Tierney et al. [64], Van der Hoek et al. [67], Van Mierlo et al. [65], Ye et al. [74].

The meta-analysis’s accuracy estimates of the sensitivity, specificity, and false positive rate, with 95% confidence intervals (CI), are presented in Table 2.

Table 2. Summary accuracy estimates—meta-analysis summary points of the sensitivity, the specificity, and the false positive rate, with 95% confidence intervals (CI).

Summary Estimates	Sensitivity (95% CI)	Specificity (95% CI)	False Positive Rate (95% CI)
Meta-analysis of all digital solutions	0.79 (0.75–0.83)	0.77 (0.73–0.81)	0.23 (0.19–0.27)
Meta-analysis of innovative digital solutions	0.82 (0.79–0.86)	0.73 (0.64–0.80)	0.27 (0.20–0.36)
Meta-analysis of paper-based digital solutions	0.77 (0.70–0.83)	0.78 (0.74–0.82)	0.22 (0.18–0.26)

For the meta-analysis of all digital solutions, a low-quality estimate of the sensitivity is 0.79 (95% CI: 0.75–0.83), and a low-quality estimate of the specificity is 0.77 (95% CI: 0.73–0.81) (Table 3). For the meta-analysis of innovative digital solutions, a moderate quality estimate of the sensitivity is 0.82 (95% IC: 0.79–0.86), and a low-quality estimate of the specificity is 0.73 (95% IC: 0.64–0.80), and for the meta-analysis of paper-based digital solutions, a low-quality estimate of the sensitivity is 0.77 (95% IC: 0.70–0.83), and a moderate quality estimate of the specificity estimate is 0.78 (95% IC: 0.74–0.82) (Table 3).

The results of each of the 24 studies in the ROC space with the quality assessment obtained using the QUADAS-2 tool, namely concerning the risk of bias and quality concerns, are present in the Supplementary Materials.

Table 3. GRADE assessment results for the meta-analysis.

Outcome	Number of Studies (Number of Participants)	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	Quality of the Evidence (GRADE)
Sensitivity (all digital solutions)	24 studies (1301 participants)	Serious (downgraded one level). Seventeen studies (71%) suggested a high risk of bias in the patient selection, and sixteen studies (67%) presented an unclear risk of bias for the index test and reference standard domains, as assessed by the QUADAS-2.	Serious (downgraded one level). High heterogeneity was identified based on the dissimilarity of the point estimates, nonoverlap of CIs through forest plots, and examining the random-effects meta-analysis figure, where the 95% prediction region is much larger than the 95% confidence region.	Not serious. Regarding indirectness, no critical differences were found between the populations studied and those for whom the recommendation is intended. Also, there is a low concern about the applicability for the index and reference test domains.	Not serious. The CI of the sensitivity summary point was considered sufficiently narrow not to demand downgrading.	None. A test for funnel plot asymmetry (Deek’s test) for assessment of publication bias suggests that the publication bias is low (p -value = 0.78).	Low
Specificity (all digital solutions)	24 studies (2069 participants)	Serious (downgraded one level). Seventeen studies (71%) suggested a high risk of bias in the patient selection, and sixteen studies (67%) presented an unclear risk of bias for the index test and reference standard domains, as assessed by the QUADAS-2.	Serious (downgraded one level). High heterogeneity was identified based on the dissimilarity of the point estimates, nonoverlap of CIs through forest plots, and examining the random-effects meta-analysis figure, where the 95% prediction region is much larger than the 95% confidence region.	Not serious. Regarding indirectness, no critical differences were found between the populations studied and those for whom the recommendation is intended. Also, there is a low concern about the applicability for the index and reference test domains.	Not serious. The CI of the specificity summary point was considered sufficiently narrow not to demand downgrading.	None. A test for funnel plot asymmetry (Deek’s test) for assessment of publication bias suggests that the publication bias is low (p -value = 0.78).	Low
Sensitivity (innovative digital solutions)	9 studies (628 participants)	Serious (downgraded one level). Six studies (67%) suggested a high risk of bias in the patient selection, eight studies (89%) presented an unclear risk of bias for the index test domain, and seven studies (78%) presented an unclear risk of bias for the reference standard domain, as assessed by the QUADAS-2.	Not serious. Low heterogeneity was identified based on the similarity of the point estimates, in the overlap of CIs with the inspection of forest plots, and in examining the random-effects meta-analysis figure, where the 95% prediction region is close to the 95% confidence region.	Not serious. Regarding indirectness, no critical differences were found between the populations studied and those for whom the recommendation is intended. Also, there is a low concern about applicability for the index and reference test domains.	Not serious. The CI of the sensitivity summary point was considered sufficiently narrow not to demand downgrading.	None. A test for funnel plot asymmetry (Deek’s test) for the assessment of publication bias suggests that the publication bias is low (p -value = 0.94).	Moderate

Table 3. Cont.

Outcome	Number of Studies (Number of Participants)	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	Quality of the Evidence (GRADE)
Specificity (innovative digital solutions)	9 studies (1079 participants)	Serious (downgraded one level). Six studies (67%) suggested a high risk of bias in the patient selection, eight studies (89%) presented an unclear risk of bias for the index test domain, and seven studies (78%) presented an unclear risk of bias for the reference standard domain, as assessed by the QUADAS-2.	Serious (downgraded one level). High heterogeneity was identified based on the dissimilarity of the point estimates, nonoverlap of CIs through forest plots, and examining the random-effects meta-analysis figure, where the 95% prediction region is much larger than the 95% confidence region.	Not serious. Regarding indirectness, no critical differences were found between the populations studied and those for whom the recommendation is intended. Also, there is a low concern about applicability for the index and reference test domains.	Not serious. The CI of the specificity summary point was considered sufficiently narrow not to demand downgrading.	None. A test for funnel plot asymmetry (Deek's test) for the assessment of publication bias suggests that the publication bias is low (p -value = 0.94).	Low
Sensitivity (paper-based digital solutions)	15 studies (673 participants)	Serious (downgraded one level). Eleven studies (73%) suggested a high risk of bias in the patient selection, eight studies (53%) presented an unclear risk of bias for the index test domain, and nine studies (60%) presented an unclear risk of bias for the reference standard domain, as assessed by the QUADAS-2.	Serious (downgraded one level). High heterogeneity was identified based on the dissimilarity of the point estimates, nonoverlap of CIs through forest plots, and examining the random-effects meta-analysis figure, where the 95% prediction region is much larger than the 95% confidence region.	Not serious. Regarding indirectness, no critical differences were found between the populations studied and those for whom the recommendation is intended. Also, there is a low concern about applicability for the index and reference test domains.	Not serious. The CI of the sensitivity summary point was considered sufficiently narrow not to demand downgrading.	None. A test for funnel plot asymmetry (Deek's test) for the assessment of publication bias suggests that the publication bias is low (p -value = 0.69).	Low
Specificity (paper-based digital solutions)	15 studies (990 participants)	Serious (downgraded one level). Eleven studies (73%) suggested a high risk of bias in the patient selection, eight studies (53%) presented an unclear risk of bias for the index test domain, and nine studies (60%) presented an unclear risk of bias for the reference standard domain, as assessed by the QUADAS-2.	Not serious. Low heterogeneity was identified based on the similarity of the point estimates, in the overlap of forest plots, and in examining the random-effects meta-analysis figure, where the 95% prediction region is close to the 95% confidence region.	Not serious. Regarding applicability, seven studies (47%) revealed low concern for the patient-selection domain, fifteen studies (100%) revealed low concern for the index test domain, and twelve studies (80%) were considered to have low concern for the reference test domain, as assessed by the QUADAS-2.	Not serious. The CI of the specificity summary point was considered sufficiently narrow not to demand downgrading.	None. A test for funnel plot asymmetry (Deek's test) for the assessment of publication bias suggests that the publication bias is low (p -value = 0.69).	Moderate

4. Discussion

This systematic review assessed the diagnostic accuracy of digital solutions used for cognitive screening, further analyzing whether these were paper-based digital solutions or innovative digital solutions. There is low- to moderate-quality evidence suggesting that digital solutions are reasonably sensitive and specific to be used for cognitive impairment screening.

The index tests assessed were quite variable, with sensitivity levels varying between 0.49 and 0.95 and specificity levels between 0.50 and 0.91. The index tests classified as innovative digital solutions offered at least a sensitivity value of 0.78 but showed lower specificity levels than the other subgroup (between 0.50 and 0.90). The index tests classified as paper-based digital solutions revealed at least a specificity value of 0.72, but sensitivity started at 0.49 (and eight studies out of fifteen reported sensitivity values below 0.78).

The study that reported higher sensitivity values among those tests classified as paper-based digital solutions reported on the Beijing version of the MoCA (sensitivity = 0.95; specificity = 0.87) [62]. This performance was similar to the MoCA paper-and-pencil version of the instrument for detecting MCI in elderly Chinese living in communities (sensitivity = 0.81; specificity = 0.83) [77], suggesting that both versions are equivalent. For the subgroup of tests classified as innovative digital solutions, the Digital Screening System [76] showed the highest sensitivity and specificity levels (sensitivity = 0.85, specificity = 0.90). These two index tests were assessed against robust reference tests (a clinical assessment performed by a team of health professionals including a neurologist, a geriatrician, and a psychiatrist (MoCA-CC [62]), and experienced doctors and neuropsychologists (Digital Screening System [76])). The assessment by a team of specialists is the gold standard for cognitive evaluation [78]. The Digital Screening System aims to assess visuospatial constructional capabilities, visual memory function, and cognitive functions, such as visuospatial abilities, visual episodic memory, organization skills, attention, and visuomotor coordination. It is based on the neuropsychological test Rey–Osterrieth Complex Figure and uses a data-driven convolutional neural network architecture through transfer learning and deep learning methods [76]. Despite being developed from inception to be applied by electronic means, most innovative digital solutions are inspired by traditional neuropsychological tests. In the innovative digital solutions subgroup, the index test Virtual Supermarket Program (VSP) stands out for using virtual reality game-based tests in screening for MCI in older adults, showing an attempt to develop a test that uses a task from daily life, potentially increasing its ecological validity. Interestingly, this test showed high sensitivity and specificity values (sensitivity = 0.85; specificity = 0.80) [73], suggesting that there is value in exploring the use of game-based tests to screen for cognitive impairment.

The index test that presented lower sensitivity and specificity in the subgroup of index tests based on paper-and-pencil tests is the MemTrax test (MTX) (sensitivity = 0.49, specificity = 0.78) [67]. This index test was based on the Continuous Recognition Task (CRT) paradigm. Among the index tests developed from inception to be applied as digital solutions, Cognivue [70] and CogEvo [58] showed the lowest sensitivity and specificity levels (Cognivue: sensitivity = 0.78, specificity = 0.50; CogEvo: sensitivity = 0.78, specificity = 0.54). These three index tests that presented the lowest sensitivity/specificity levels were compared against a reference standard consisting of only brief cognitive screening instruments (i.e., MoCA, SLUMS, and MMSE, respectively). The MoCA paper-and-pencil test demonstrated a sensitivity of 90% and a specificity of 87% for detecting MCI [15]. The MMSE paper-and-pencil test showed a pooled sensitivity of 85% and a specificity of 86% in a non-clinical community setting [79]. The SLUMS paper-and-pencil test for detecting MCI in patients with less than a high school education had a sensitivity of 92% and a specificity of 81%, and in patients with a high school education or more, a sensitivity of 95% and a specificity of 76% [17]. Despite the relatively high sensitivity and specificity levels, these instruments are not the gold standard for cognitive assessment and, therefore, their use might have affected the sensitivity and specificity calculations of the index test and, certainly, undermines the confidence in the reported results.

The early detection of cognitive impairment is critical to an early intervention [12,13]. Index tests with high sensitivity levels are essential when the goal is to identify a serious disease with available treatment [44,80]. Digital solutions emerge as a valid alternative for cognitive screening, potentially enhancing cognitive screening and monitoring in the general and clinical population, since most do not require the presence of a trained professional and have an automatic digital screening system and scoring [52,62,76], decreasing the costs associated with their use and facilitating the screening for high numbers of individuals. Digital solutions can be valuable in neuropsychological assessment, enabling the development of large-scale, norm-based, and technology-driven tests [28]. These tools produce large cognitive datasets that can be informative through machine learning and big data analysis, contributing to the detection of patterns and declines in cognitive performance [28]. The accuracy estimates of sensitivity, specificity, and the false positive rate found in this meta-analysis suggest that digital solutions have satisfactory accuracy and the potential to be used as instruments for cognitive screening. However, these estimates must be interpreted and compared with caution, since the GRADE evidence assessment and rating of these accuracy estimates mainly showed a low quality of evidence. The risk of bias and inconsistency found in the GRADE assessment downgraded the quality of the evidence.

The quality of the included studies as evaluated by the QUADAS-2 tool suggests a risk of bias in the patients' selection domain, including for those studies presenting the digital index tests with higher sensitivity/specificity values. A test accuracy study with a high risk of bias in the participant selection domain can give inflated estimates of sensitivity and specificity [81]. Despite the different definitions used by the studies, we found relative homogeneity in the target condition, as they all focus on the diagnostic ability and accuracy when screening for cognitive impairment. Nevertheless, the reference standards display substantial methodological heterogeneity. This heterogeneity was due to significant variations in the instruments adopted and/or the clinical assessment process followed across the studies. A similar reference standard, preferentially a gold standard, should be applied across studies to facilitate accuracy comparisons and increase the confidence in the results [39].

Considering the heterogeneity in reference standards and index tests across studies, the meta-analysis estimates have limitations, and the interpretation and comparison of estimates should be performed cautiously [43]. Also, the high risk of bias in the patients' selection downgraded the quality of the evidence. When applying the GRADE approach, overall, there was a serious risk of bias due to less robust procedures regarding patient selection, the index test, and reference standard domains, and consequently, the quality of evidence was downgraded by one level in this domain. The high heterogeneity of the outcomes of the included studies also prompted the downgrading of the evidence due to inconsistency by one level. These aspects must be considered in the design of cognitive diagnostic accuracy studies to improve the quality of the evidence.

Future studies should adopt more rigorous, at-random sampling procedures to reduce the probability of the risk of bias from patient recruitment. Also, future diagnostic tools should consider adopting a similar gold standard to a reference test to facilitate comparisons and increase the confidence in the results. Gold standards involve the assessment of multiple cognitive domains, including memory, by qualified professionals [78]. However, investigators and practitioners must consider the diagnostic properties of the different digital solutions and the reference test against which the accuracy values were calculated to make an informed choice.

The impact of participants' digital skills on the access to and on the results of digitally administered tests can also be addressed in future studies. Also, the feasibility of digital solutions for cognitive remote screening for specific populations also needs to be investigated in future studies.

5. Conclusions

There is low- to moderate-quality evidence that digital solutions can be used for cognitive screening, but more high-quality research is needed. A careful assessment of the accuracy levels and quality of the evidence of each digital solution is recommended before considering its use.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/app14062640/s1>. Table S1: Characteristics of the included studies; Figure S1: Random-effects meta-analysis—Summary ROC (SROC) curve with quality assessment obtained by the QUADAS-2 tool; Figure S2: Random-effects meta-analysis—Summary ROC (SROC) curve with quality assessment obtained by the QUADAS-2 tool (applicability concerns).

Author Contributions: Conceptualization, A.I.M., J.P., A.G.S. and N.P.R.; methodology, M.M., A.I.M., J.P., A.G.S. and N.P.R.; software, M.M.; validation, M.M., A.G.S. and N.P.R.; formal analysis, M.M., A.I.M., A.G.S. and N.P.R.; investigation, M.M., A.I.M., A.G.S. and N.P.R.; data curation, M.M. and N.P.R.; writing—original draft preparation, M.M. and N.P.R.; writing—review and editing, M.M., A.G.S. and N.P.R.; visualization, M.M.; supervision, A.G.S. and N.P.R.; project administration, N.P.R. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the PRR—Plano de Recuperação e Resiliência, and by the NextGenerationEU funds at the University of Aveiro, through the scope of the Agenda for Business Innovation “HfPT—Health from Portugal” (Project no. 41 with the application C644937233-00000047).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data needed to evaluate the conclusions are present in the paper. Additional data related to this paper are available upon request from the corresponding author, M.M.

Conflicts of Interest: Author Joana Pais was employed by the company Neuroinova, Lda. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Holsinger, T.; Deveau, J.; Boustani, M.; Williams, J.W. Does this patient have dementia? *JAMA* **2007**, *297*, 2391–2404. [[CrossRef](#)] [[PubMed](#)]
2. Plassman, B.L.; Williams, J.W.J.; Burke, J.R.; Holsinger, T.; Benjamin, S. Systematic review: Factors associated with risk for and possible prevention of cognitive decline in later life. *Ann. Intern. Med.* **2010**, *153*, 182–193. [[CrossRef](#)] [[PubMed](#)]
3. Ritchie, K.; Carriere, I.; Ritchie, C.W.; Berr, C.; Artero, S.; Ancelin, M.-L. Designing prevention programmes to reduce incidence of dementia: Prospective cohort study of modifiable risk factors. *BMJ* **2010**, *341*, c3885. [[CrossRef](#)] [[PubMed](#)]
4. Livingston, G.; Sommerlad, A.; Orgeta, V.; Costafreda, S.G.; Huntley, J.; Ames, D.; Ballard, C.; Banerjee, S.; Burns, A.; Cohen-Mansfield, J.; et al. Dementia prevention, intervention, and care. *Lancet* **2017**, *390*, 2673–2734. [[CrossRef](#)]
5. Boyle, P.A.; Buchman, A.S.; Wilson, R.S.; Leurgans, S.E.; Bennett, D.A. Physical frailty is associated with incident Mild Cognitive Impairment in community-based older persons. *J. Am. Geriatr. Soc.* **2010**, *58*, 248–255. [[CrossRef](#)] [[PubMed](#)]
6. US Preventive Services Task Force. Screening for cognitive impairment in older adults: US Preventive Services Task Force recommendation statement. *JAMA* **2020**, *323*, 757–763. [[CrossRef](#)]
7. Sabbagh, M.N.; Boada, M.; Borson, S.; Chilukuri, M.; Dubois, B.; Ingram, J.; Iwata, A.; Porsteinsson, A.P.; Possin, K.L.; Rabinovici, G.D.; et al. Early detection of Mild Cognitive Impairment (MCI) in primary care. *J. Prev. Alzheimer's Dis.* **2020**, *7*, 165–170. [[CrossRef](#)]
8. Jack, C.R.; Bennett, D.A.; Blennow, K.; Carrillo, M.C.; Dunn, B.; Haeblerlein, S.B.; Holtzman, D.M.; Jagust, W.; Jessen, F.; Karlawish, J.; et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's Dement.* **2018**, *14*, 535–562. [[CrossRef](#)]
9. Petersen, R.C.; Caracciolo, B.; Brayne, C.; Gauthier, S.; Jelic, V.; Fratiglioni, L. Mild Cognitive Impairment: A concept in evolution. *J. Intern. Med.* **2014**, *275*, 214–228. [[CrossRef](#)]
10. Roberts, R.O.; Knopman, D.S.; Mielke, M.M.; Cha, R.H.; Pankratz, V.S.; Christianson, T.J.; Geda, Y.E.; Boeve, B.F.; Ivnik, R.J.; Tangalos, E.G.; et al. Higher risk of progression to dementia in Mild Cognitive Impairment cases who revert to normal. *Neurology* **2014**, *82*, 317–325. [[CrossRef](#)]
11. Limpawattana, P.; Manjavong, M. The Mini-Cog, Clock Drawing Test, and Three-Item Recall Test: Rapid cognitive screening tools with comparable performance in detecting Mild NCD in older patients. *Geriatrics* **2021**, *6*, 91. [[CrossRef](#)]

12. De Roeck, E.E.; De Deyn, P.P.; Dierckx, E.; Engelborghs, S. Brief cognitive screening instruments for early detection of Alzheimer's disease: A systematic review. *Alzheimer's Res. Ther.* **2019**, *11*, 21. [\[CrossRef\]](#)
13. Brodaty, H.; Low, L.-F.; Gibson, L.; Burns, K. What is the best dementia screening instrument for general practitioners to use? *Am. J. Geriatr. Psychiatry* **2006**, *14*, 391–400. [\[CrossRef\]](#)
14. Folstein, M.F.; Folstein, S.E.; McHugh, P.R. "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **1975**, *12*, 189–198. [\[CrossRef\]](#)
15. Nasreddine, Z.S.; Phillips, N.A.; Bédirian, V.; Charbonneau, S.; Whitehead, V.; Collin, I.; Cummings, J.L.; Chertkow, H. The Montreal Cognitive Assessment, MoCA: A brief screening tool for Mild Cognitive Impairment. *J. Am. Geriatr. Soc.* **2005**, *53*, 695–699. [\[CrossRef\]](#)
16. Walterfang, M.; Siu, R.; Velakoulis, D. The NUCOG: Validity and reliability of a brief cognitive screening tool in neuropsychiatric patients. *Aust. N. Z. J. Psychiatry* **2006**, *40*, 995–1002. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Tariq, S.H.; Tumosa, N.; Chibnall, J.T.; Perry, M.H., 3rd; Morley, J.E. Comparison of the Saint Louis University Mental Status examination and the Mini-Mental State Examination for detecting dementia and mild neurocognitive disorder—A pilot study. *Am. J. Geriatr. Psychiatry* **2006**, *14*, 900–910. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Scharre, D.W.; Chang, S.I.; Murden, R.A.; Lamb, J.; Beversdorf, D.Q.; Kataki, M.; Nagaraja, H.N.; Bornstein, R.A. Self-Administered Gerocognitive Examination (SAGE): A brief cognitive assessment instrument for Mild Cognitive Impairment (MCI) and early dementia. *Alzheimer Dis. Assoc. Disord.* **2010**, *24*, 64–71. [\[CrossRef\]](#)
19. Elamin, M.; Holloway, G.; Bak, T.H.; Pal, S. The utility of the Addenbrooke's Cognitive Examination Version Three in early-onset dementia. *Dement. Geriatr. Cogn. Disord.* **2016**, *41*, 9–15. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Dsurney, J. Alzheimer's Quick Test: Assessment of parietal lobe function. *Appl. Neuropsychol.* **2007**, *14*, 232–233. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Takechi, H.; Dodge, H.H. Scenery Picture Memory Test: A new type of quick and effective screening test to detect early stage Alzheimer's disease patients. *Geriatr. Gerontol. Int.* **2010**, *10*, 183–190. [\[CrossRef\]](#)
22. Buschke, H.; Kuslansky, G.; Katz, M.; Stewart, W.; Sliwinski, M.; Eckholdt, H.; Lipton, R. Screening for dementia with the Memory Impairment Screen. *Neurology* **1999**, *52*, 231. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Borson, S.; Scanlan, J.M.; Chen, P.; Ganguli, M. The Mini-Cog as a screen for dementia: Validation in a population-based sample. *J. Am. Geriatr. Soc.* **2003**, *51*, 1451–1454. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Borson, S.; Scanlan, J.; Brush, M.; Vitaliano, P.; Dokmak, A. The Mini-Cog: A cognitive 'vital signs' measure for dementia screening in multi-lingual elderly. *Int. J. Geriatr. Psychiatry* **2000**, *15*, 1021–1027. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Ehreke, L.; Luppá, M.; König, H.-H.; Riedel-Heller, S.G. Is the Clock Drawing Test a screening tool for the diagnosis of Mild Cognitive Impairment? A systematic review. *Int. Psychogeriatr.* **2010**, *22*, 56–63. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Siqueira, G.S.A.; Hagemann, P.D.M.; Coelho, D.D.S.; Santos, F.H.D.; Bertolucci, P.H. Can MoCA and MMSE be interchangeable cognitive screening tools? A systematic review. *Gerontologist* **2019**, *59*, e743–e763. [\[CrossRef\]](#)
27. Sabbagh, M.N.; Boada, M.; Borson, S.; Doraiswamy, P.M.; Dubois, B.; Ingram, J.; Iwata, A.; Porsteinsson, A.P.; Possin, K.L.; Rabinovici, G.D.; et al. Early detection of Mild Cognitive Impairment (MCI) in an at-home setting. *J. Prev. Alzheimer's Dis.* **2020**, *7*, 171–178. [\[CrossRef\]](#)
28. Diaz-orueta, U.; Blanco-Campal, A.; Lamar, M.; Libon, D.J.; Burke, T. Marrying past and present neuropsychology: Is the future of the process-based approach technology-based? *Front. Psychol.* **2020**, *11*, 361. [\[CrossRef\]](#)
29. Berg, J.-L.; Durant, J.; Léger, G.C.; Cummings, J.L.; Nasreddine, Z.; Miller, J.B. Comparing the electronic and standard versions of the Montreal Cognitive Assessment in an outpatient memory disorders clinic: A validation study. *J. Alzheimer's Dis.* **2018**, *62*, 93–97. [\[CrossRef\]](#)
30. Wohlfahrt-Laymann, J.; Hermens, H.; Villalonga, C.; Vollenbroek-Hutten, M.; Banos, O. MobileCognitracker. *J. Ambient Intell. Humaniz. Comput.* **2019**, *10*, 2143–2160. [\[CrossRef\]](#)
31. Lauraitis, A.; Maskeliūnas, R.; Damaševičius, R.; Krilavičius, T. A mobile application for smart computer-aided self-administered testing of cognition, speech, and motor impairment. *Sensors* **2020**, *20*, 3236. [\[CrossRef\]](#)
32. Hansen, T.I.; Hafnerstrom, E.C.D.; Brunner, J.F.; Lehn, H.; Håberg, A.K. Initial validation of a web-based self-administered neuropsychological test battery for older adults and seniors. *J. Clin. Exp. Neuropsychol.* **2015**, *37*, 581–594. [\[CrossRef\]](#)
33. Assmann, K.E.; Bailet, M.; Lecoffre, A.C.; Galan, P.; Herberg, S.; Amieva, H.; Kesse-Guyot, E. Comparison between a self-administered and supervised version of a web-based cognitive test battery: Results from the NutriNet-Santé cohort study. *J. Med. Internet Res.* **2016**, *18*, e68. [\[CrossRef\]](#)
34. Morris, R.G.; Evenden, J.L.; Sahakian, B.J.; Robbins, T.W. Computer-aided assessment of dementia: Comparative studies of neuropsychological deficits in Alzheimer-type dementia and Parkinson's disease. In *Cognitive Neurochemistry*; Oxford University Press: Oxford, UK, 1987; pp. 21–36.
35. Bevilacqua, R.; Barbarossa, F.; Fantechi, L.; Fornarelli, D.; Paci, E.; Bolognini, S.; Giammarchi, C.; Lattanzio, F.; Paciaroni, L.; Riccardi, G.R.; et al. Radiomics and artificial intelligence for the diagnosis and monitoring of Alzheimer's disease: A systematic review of studies in the field. *J. Clin. Med.* **2023**, *12*, 5432. [\[CrossRef\]](#)
36. Bevilacqua, R.; Felici, E.; Cucchieri, G.; Amabili, G.; Margaritini, A.; Franceschetti, C.; Barboni, I.; Paolini, S.; Civerchia, P.; Raccichini, A.; et al. Results of the italian RESILIEN-T pilot study: A mobile health tool to support older people with Mild Cognitive Impairment. *J. Clin. Med.* **2023**, *12*, 6129. [\[CrossRef\]](#)

37. Pereira, C.R.; Pereira, D.R.; Weber, S.A.; Hook, C.; de Albuquerque, V.H.C.; Papa, J.P. A survey on computer-assisted Parkinson's disease diagnosis. *Artif. Intell. Med.* **2019**, *95*, 48–63. [[CrossRef](#)] [[PubMed](#)]
38. Lumsden, J.; Edwards, E.A.; Lawrence, N.S.; Coyle, D.; Munafò, M.R. Gamification of cognitive assessment and cognitive training: A systematic review of applications and efficacy. *JMIR Serious Games* **2016**, *4*, e11. [[CrossRef](#)] [[PubMed](#)]
39. Macaskill, P.; Gatsonis, C.; Deeks, J.; Harbord, R.; Takwoingi, Y. Analysing and presenting results. In *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*; Cochrane: London, UK, 2010; pp. 1–61.
40. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; Group, T.P. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA statement. *PLoS Med.* **2009**, *6*, e1000097. [[CrossRef](#)] [[PubMed](#)]
41. Silva, A.; Rocha, N.; Martins, A.; Pais, J. Diagnostic accuracy of digital solutions to screen for cognitive impairment: A systematic review. Available online: <http://srdta.cochrane.org/> (accessed on 31 March 2022).
42. Cohen, J.F.; Korevaar, D.A.; Altman, D.G.; Bruns, D.E.; Gatsonis, C.A.; Hooft, L.; Irwig, L.; Levine, D.; Reitsma, J.B.; de Vet, H.C.W.; et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: Explanation and elaboration. *BMJ Open* **2016**, *6*, e012799. [[CrossRef](#)]
43. Macaskill, P.; Takwoingi, Y.; Deeks, J.J.; Gatsonis, C. Understanding meta-analysis. In *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2023; pp. 203–247. [[CrossRef](#)]
44. Ghaaliq, A.; Mb, L.; Frca, C.; Mccluskey, A.; Chb, M.B. Clinical tests: Sensitivity and specificity. *Contin. Educ. Anaesth. Crit. Care Pain* **2008**, *8*, 221–223.
45. Whiting, P.F.; Rutjes, A.W.S.; Westwood, M.E.; Mallett, S.; Deeks, J.J.; Reitsma, J.B.; Leeflang, M.M.G.; Sterne, J.A.C.; Bossuyt, P.M.M.; QUADAS-2 Group. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann. Intern. Med.* **2011**, *155*, 529–536. [[CrossRef](#)]
46. Schünemann, H.J.; Mustafa, R.A.; Brozek, J.; Steingart, K.R.; Leeflang, M.; Murad, M.H.; Bossuyt, P.; Glasziou, P.; Jaeschke, R.; Lange, S.; et al. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. *J. Clin. Epidemiol.* **2020**, *122*, 129–141. [[CrossRef](#)]
47. Schünemann, H.J.; Mustafa, R.A.; Brozek, J.; Steingart, K.R.; Leeflang, M.; Murad, M.H.; Bossuyt, P.; Glasziou, P.; Jaeschke, R.; Lange, S.; et al. GRADE guidelines: 21 part 2. Test accuracy: Inconsistency, imprecision, publication bias, and other domains for rating the certainty of evidence and presenting it in evidence profiles and summary of findings tables. *J. Clin. Epidemiol.* **2020**, *122*, 142–152. [[CrossRef](#)]
48. Dinnes, J.; Deeks, J.J.; Leeflang, M.M.; Li, T. Collecting data. In *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2023; pp. 131–167. [[CrossRef](#)]
49. Takwoingi, Y.; Dendukuri, N.; Schiller, I.; Rücker, G.; Jones, H.E.; Partlett, C.; Macaskill, P. Undertaking meta-analysis. In *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2023; pp. 249–325. [[CrossRef](#)]
50. Patel, A.; Cooper, N.; Freeman, S. Graphical enhancements to summary receiver operating characteristic plots to facilitate the analysis and reporting of meta-analysis of diagnostic test accuracy data. *Res. Synth. Methods* **2021**, *12*, 34–44. [[CrossRef](#)] [[PubMed](#)]
51. Freeman, S.C.; Kerby, C.R.; Patel, A.; Cooper, N.J.; Quinn, T.; Sutton, A.J. Development of an interactive web-based tool to conduct and interrogate meta-analysis of diagnostic test accuracy studies: MetaDTA. *BMC Med. Res. Methodol.* **2019**, *19*, 81. [[CrossRef](#)]
52. Kokubo, N.; Yokoi, Y.; Saitoh, Y.; Murata, M.; Maruo, K.; Takebayashi, Y.; Shinmei, I.; Yoshimoto, S.; Horikoshi, M. A new device-aided cognitive function test, User eXperience-Trail Making Test (UX-TMT), sensitively detects neuropsychological performance in patients with dementia and Parkinson's disease. *BMC Psychiatry* **2018**, *18*, 220. [[CrossRef](#)]
53. Rhodius-meester, H.F.M.; PaaJanen, T.; Koikkalainen, J.; Mahdiani, S.; Bruun, M.; Baroni, M.; Lemstra, A.W.; Scheltens, P.; Herukka, S.; Pikkariainen, M.; et al. cCOG: A web-based cognitive test tool for detecting neurodegenerative disorders. *Alzheimer's Dement.* **2020**, *12*, e12083. [[CrossRef](#)]
54. Brandt, J.; Blehar, J.; Anderson, A.; Gross, A.L. Further validation of the Internet-based Dementia Risk Assessment. *J. Alzheimer's Dis.* **2014**, *41*, 937–945. [[CrossRef](#)] [[PubMed](#)]
55. Kalafatis, C.; Modarres, M.H.; Apostolou, P.; Marefat, H. Validity and cultural generalisability of a 5-minute AI-based, computerised cognitive assessment in Mild Cognitive Impairment and Alzheimer's dementia. *Front. Psychiatry* **2021**, *12*, 706695. [[CrossRef](#)]
56. Curiel, R.E.; Crocco, E.; Rosado, M.; Duara, R.; Greig, M.T. A brief computerized paired associate test for the detection of Mild Cognitive Impairment in community-dwelling older adults. *J. Alzheimer's Dis.* **2016**, *54*, 793–799. [[CrossRef](#)] [[PubMed](#)]
57. Alegret, M.; Muñoz, N.; Roberto, N.; Rentz, D.M.; Valero, S.; Gil, S.; Marquí, M.; Hernández, I.; Riveros, C.; Sanabria, A.; et al. A computerized version of the Short Form of the Face-Name Associative Memory Exam (FACEmemory[®]) for the early detection of Alzheimer's disease. *Alzheimer's Res. Ther.* **2020**, *12*, 25. [[CrossRef](#)] [[PubMed](#)]
58. Ichii, S.; Nakamura, T.; Kawarabayashi, T.; Takatama, M.; Ohgami, T.; Ihara, K.; Shoji, M. CogEvo, a cognitive function balancer, is a sensitive and easy psychiatric test battery for age-related cognitive decline. *Geriatr. Gerontol. Int.* **2019**, *20*, 248–255. [[CrossRef](#)] [[PubMed](#)]
59. Wong, A.; Fong, C.; Mok, V.C.; Leung, K. Computerized Cognitive Screen (CoCoSc): A self-administered computerized test for screening for cognitive impairment in community social centers. *J. Alzheimer's Dis.* **2017**, *59*, 1299–1306. [[CrossRef](#)] [[PubMed](#)]

60. Buckley, R.F.; Sparks, K.; Papp, K.; Dekhtyar, M.; Martin, C.; Burnham, S.; Sperling, R.; Rentz, D. Computerized cognitive testing for use in clinical trials: A comparison of the NIH Toolbox and Cogstate C3 Batteries. *J. Prev. Alzheimer's Dis.* **2017**, *4*, 3–11. [[CrossRef](#)] [[PubMed](#)]
61. Memória, C.M.; Yassuda, M.S.; Nakano, E.Y.; Forlenza, O.V. Contributions of the Computer-Administered Neuropsychological Screen for Mild Cognitive Impairment (CANS-MCI) for the diagnosis of MCI in Brazil. *Int. Psychogeriatr.* **2014**, *26*, 1483–1491. [[CrossRef](#)] [[PubMed](#)]
62. Yu, K.; Zhang, S.; Wang, Q.; Wang, X.; Qin, Y.; Wang, J.; Li, C.; Wu, Y.; Wang, W.; Lin, H. Development of a computerized tool for the chinese version of the Montreal Cognitive Assessment for screening Mild Cognitive Impairment. *Int. Psychogeriatr.* **2015**, *27*, 213–219. [[CrossRef](#)]
63. Scharre, D.W.; Chang, S.; Nagaraja, H.N.; Vrettos, N.E.; Bornstein, R.A. Digitally translated Self-Administered Gerocognitive Examination (eSAGE): Relationship with its validated paper version, neuropsychological evaluations, and clinical assessments. *Alzheimer's Res. Ther.* **2017**, *9*, 44. [[CrossRef](#)]
64. Tierney, M.C.; Naglie, G.; Upshur, R.; Moineddin, R.; Charles, J.; Jaakkimainen, R.L. Feasibility and validity of the self-administered computerized assessment of Mild Cognitive Impairment with older primary care patients. *Alzheimer Dis. Assoc. Disord.* **2014**, *28*, 311–319. [[CrossRef](#)]
65. Van Mierlo, L.D.; Wouters, H.; Sikkes, S.A.M.; Van Der Flier, W.M. Screening for Mild Cognitive Impairment and dementia with automated, anonymous online and Telephone Cognitive Self-Tests. *J. Alzheimer's Dis.* **2017**, *56*, 249–259. [[CrossRef](#)]
66. Dougherty, J.H.; Cannon, R.L.; Nicholas, C.R.; Hall, L.; Hare, F. The Computerized Self Test (CST): An interactive, internet accessible cognitive screening test for dementia. *J. Alzheimer's Dis.* **2010**, *20*, 185–195. [[CrossRef](#)]
67. Van Der Hoek, M.D.; Nieuwenhuizen, A.; Keijer, J.; Ashford, J.W. The MemTrax Test compared to the Montreal Cognitive Assessment estimation of Mild Cognitive Impairment. *J. Alzheimer's Dis.* **2019**, *67*, 1045–1054. [[CrossRef](#)] [[PubMed](#)]
68. Scanlon, L.; O'Shea, E.; O'Caomh, R.; Timmons, S. Usability and validity of a battery of computerised cognitive screening tests for detecting cognitive impairment. *Gerontology* **2015**, *62*, 247–252. [[CrossRef](#)] [[PubMed](#)]
69. Fung, A.W.; Chiu, L.; Lam, W.; Fung, A.W. Validation of a computerized Hong Kong-vigilance and memory test (HK-VMT) to detect early cognitive impairment in healthy older adults. *Aging Ment. Health* **2020**, *24*, 185–191. [[CrossRef](#)] [[PubMed](#)]
70. Cahn-hidalgo, D.; Estes, P.W.; Benabou, R. Validity, reliability, and psychometric properties of a computerized, cognitive assessment test (Cognivue®). *World J. Psychiatry* **2020**, *10*, 1–11. [[CrossRef](#)]
71. Rodríguez-Salgado, A.M.; Llibre-Guerra, J.J.; Tsoy, E.; Peñalver-Guia, A.I.; Bringas, G.; Erhoff, S.J.; Kramer, J.H.; Allen, I.E.; Valcour, V.; Miller, B.L.; et al. A brief digital cognitive assessment for detection of cognitive impairment in cuban older adults. *J. Alzheimer's Dis.* **2021**, *79*, 85–94. [[CrossRef](#)]
72. Liu, X.; Chen, X.; Zhou, X.; Shang, Y.; Xu, F.; Zhang, J.; He, J.; Zhao, F.; Du, B.; Wang, X.; et al. Validity of the MemTrax Memory Test compared to the Montreal Cognitive Assessment in the detection of Mild Cognitive Impairment and dementia due to Alzheimer's disease in a chinese cohort. *J. Alzheimer's Dis.* **2021**, *80*, 1257–1267. [[CrossRef](#)]
73. Yan, M.; Yin, H.; Meng, Q.; Wang, S.; Ding, Y.; Li, G.; Wang, C.; Chen, L. A Virtual Supermarket Program for the screening of Mild Cognitive Impairment in older adults: Diagnostic accuracy study. *JMIR Serious Games* **2021**, *9*, e30919. [[CrossRef](#)]
74. Ye, S.; Sun, K.; Huynh, D.; Phi, H.Q.; Ko, B.; Huang, B.; Ghomi, R.H. A computerized cognitive test battery for detection of dementia and Mild Cognitive Impairment: Instrument validation study. *JMIR Aging* **2022**, *5*, e36825. [[CrossRef](#)]
75. Paterson, T.S.E.; Sivajohan, B.; Gardner, S.; Binns, M.A.; Stokes, K.A.; Freedman, M.; Levine, B.; Troyer, A.K. Accuracy of a self-administered online cognitive assessment in detecting Amnesic Mild Cognitive Impairment. *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.* **2022**, *77*, 341–350. [[CrossRef](#)]
76. Cheah, W.; Hwang, J.; Hong, S.; Fu, L.; Chang, Y. A digital screening system for Alzheimer disease based on a neuropsychological test and a convolutional neural network: System development and validation. *JMIR Med. Inform.* **2022**, *10*, e31106. [[CrossRef](#)]
77. Lu, J.; Li, D.; Li, F.; Zhou, A.; Wang, F.; Zuo, X.; Jia, X.-F.; Song, H.; Jia, J. Montreal Cognitive Assessment in detecting cognitive impairment in chinese elderly individuals: A population-based study. *J. Geriatr. Psychiatry Neurol.* **2011**, *24*, 184–190. [[CrossRef](#)] [[PubMed](#)]
78. Petersen, R.C.; Stevens, J.; Ganguli, M.; Tangalos, E.G.; Cummings, J.; DeKosky, S.T. Practice parameter: Early detection of dementia: Mild Cognitive Impairment (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology* **2001**, *56*, 1133–1142. [[CrossRef](#)] [[PubMed](#)]
79. Mitchell, A.J. A meta-analysis of the accuracy of the Mini-Mental State Examination in the detection of dementia and Mild Cognitive Impairment. *J. Psychiatr. Res.* **2009**, *43*, 411–431. [[CrossRef](#)] [[PubMed](#)]
80. World Health Organization. *WHO Handbook for Guideline Development*; World Health Organization: Geneva, Switzerland, 2014.
81. Reitsma, J.B.; Rutjes, A.W.; Whiting, P.; Yang, B.; Leeflang, M.M.; Bossuyt, P.M.; Deeks, J.J. Assessing risk of bias and applicability. In *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2023; pp. 169–201. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.